

ANOOSHKA BAJAJ

✉ anooshkabajaj@gmail.com
🌐 linkedin.com/in/anooshkabajaj

☎ (812) 822-7876
🐙 github.com/ab490

📍 Bloomington, IN (open to relocation)
🌐 ab490.github.io

EDUCATION

Master of Science (M.S.) in Data Science

Indiana University Bloomington, IN, US

Aug 2024 – May 2026

- **GPA:** 3.90 / 4.00
- **Research Assistant (R.A.):** Working in the Cognitive AI lab, focusing on in-context learning mechanisms in Transformers. Investigating LLMs interpretability and correlating findings with human brain language understanding.
- **Relevant Courses:** Introduction to Statistics, Applied Machine Learning, Computer Vision, Advanced Database Concepts.

Bachelor of Technology (B.Tech.) in Bio-Engineering with Minor in Computer Science Engineering

Indian Institute of Technology (IIT) Mandi, India

Aug 2019 – May 2023

- **CGPA:** 8.50 / 10.00
- **Relevant Courses:** Data Science I (Data Analysis), Data Science II (Probability and Statistics), Data Science III (Machine Learning), Deep Learning and Applications, Biostatistics, Data Handling and Visualization, Data Structures and Algorithms.

EXPERIENCE

Software Engineer - Radisys India Ltd.

Bangalore, Karnataka, India

Jul 2023 – Jun 2024

(Intern : Jul 2023 – Nov 2023)

- Improved **5G** network infrastructure by collaborating across teams and optimizing OAM codebase in **C++** and **Python**.
- Managed feature requests across multiple modules, including Configuration Management (**CM**) (added YANG model configurations as per 3GPP specifications) and Performance Management (**PM**) (handled performance counters).
- Increased operational efficiency and addressed client requests related to **CU** (Centralized Unit) by presenting solutions and managing code coordination across 3 teams using **BitBucket**, ensuring clear communication and integration.

Data Engineer Intern - Technocolabs Softwares Inc.

Indore, Madhya Pradesh, India (Remote)

Jul 2022– Sep 2022

- Pre-processed dataset of over 1 million rows in **Python**, performing Exploratory Data Analysis and Feature Engineering.
- Built data warehouse with **BigQuery** leveraging **Cloud Computing**. Managed ETL workflows, generated DAGs using **GCP (Google Cloud Composer)** and developed interactive dashboards on Google Data Studio with 20+ visualizations.

Research Intern – IISER Bhopal

Bhopal, Madhya Pradesh, India

Jun 2022– Jul 2022

- Assisted Dr. Sunando Datta in analyzing 2,212 YAP target genes in the Hippo Signaling Pathway to inhibit tumor growth.
- Produced the dataset by **web scraping** and performed **data visualization** of target genes belonging to 28 gene groups.
- Deployed a **Python** web app for cell migration analysis from time-series coordinate files using **Flask**.

PROJECTS

Skin Cancer Classification App | Deep Learning, CNN, Flask

- Developed a **web application** to predict and classify skin cancer types from dermatoscopic images uploaded by users, utilizing **Convolutional Neural Networks (CNNs)** for image classification.
- Trained the deep learning model on HAM10000 dataset to classify skin cancer into one of 7 categories.
- Enhanced model accuracy by evaluating over different architectures and implementing three regularization techniques, resulting in a final deep learning model integrated into an efficient **Flask**-based web application.

Events Booking Web Application | NoSQL, MongoDB, PowerBI, React, Node.js

- Led a team of three to design and deploy a full-stack **MERN web application** for event booking, featuring role-based access control for organizers to list events and attendees to book tickets seamlessly.
- Utilized **MongoDB** as the primary database with indexed schemas across three collections to optimize data retrieval through efficient **NoSQL** document structuring and integrated it with **Power BI** for real-time data visualization.
- Built React interface with 4 components and analytics dashboards, powered by Node.js backend, hosted on **Render**.

Audio Summarizer App | GenAI, NLP, LLM, Streamlit

- Developed an audio summarization **web application** using OpenAI's **Whisper** model for transcription and Hugging Face's **Transformers (t5-small)** for generating concise summaries.
- Built a **Streamlit**-based interface with PyTorch integration for efficient processing of WAV/MP3 audio files up to 200MB.

TECHNICAL SKILLS

Programming Languages: C++, Python, R, MATLAB

Development Tools & IDEs: Jupyter, Conda, Visual Studio Code, Spyder, RStudio

Databases & Data Management: SQL, NoSQL, MySQL, PostgreSQL, MongoDB, Tableau, Power BI, MS Excel

Libraries & Frameworks: NumPy, Pandas, Matplotlib, Scikit-Learn, TensorFlow, PyTorch, Flask, Git

Certifications: **Machine Learning:** Kaggle, **Introduction to Snowflake:** DataCamp, **MATLAB Onramp:** Mathworks